Factors that Correlate with the U.S. Medical Licensure Examination Step-2 Scores in a Diverse Medical Student Population

Dotun Ogunyemi, MD and De Shawn Taylor-Harris, MD
Los Angeles, CA

Presented at the Council on Resident Education in Obstetrics and Gynecology/Association of Professors of Gynecology and Obstetrics (CREOG/APGO) Annual Conference 2004

Objective: To assess factors that correlate with performance on U.S. Medical Licensure Examination (USMLE) Step-2 examination. Our hypothesis was that demographic factors, faculty assessments and other standardized test scores will correlate with students' performance on USMLE Step 2.

Study Design: A comparison of standardized examinations and demographic factors with USMLE Step-2 scores as the outcome variable was accomplished using the educational records of 171 medical students.

Results: Mean USMLE Step 2, USMLE Step 1, NBME-OB/GYN and MCAT scores, respectively, were 190.63, 194.53, 67.47 and 24.03. Positive correlations of USMLE Step 2 were: USMLE Step-1 scores (r=0.681, p=0. 000); MCAT scores (r=0.524, p=0.000); NBME-OB/GYN scores (r=0.614, p=0.000); year of OB/GYN rotation (r=0.432, p=0.000); faculty grades (r=0.400, p=0.000); undergraduate GPA (r=0.287, p=0.000); and science GPA (r=0.255, p=0.002). Negative correlations of USMLE Step 2 were students' increasing age (r=-0.405, p=0.000), increasing number of MCAT attempts (r=-0.182, p=0.000) and increasing number of NBME-OB/GYN attempts (r=-0.310, p=0.000). There was no correlation with gender or race. Logistic regression analysis showed that a failing NBME-OB/GYN score (p=0.008), failing USMLE Step-1 score (p=0.01), failing faculty grade (p=0.029) and multiple MCAT attempts (p=0.033) independently increased the risk of failing USMLE Step 2.

Conclusions: Premedical test-taking abilities (MCAT scores), results of preclinical standardized tests (USMLE Step 1) and performance on clinical rotations (NBME-OB/GYN score and faculty grade) correlate significantly with USMLE scores. These findings may assist educators in selecting medical students at risk of performing poorly on the USMLE Step-2 examinations.

Key words: medical education ■ standardized examinations ■ clinical clerkships ■ minority medical school

 From Department of Obstetrics and Gynecology, Charles R. Drew College of Medicine and Science, Los Angeles, CA. Send correspondence and reprint requests for *J Natl Med Assoc.* 2005;97:1258–1262 to: Dotun Ogunyemi, MD, Department of Obstetrics and Gynecology, King Drew Medical Center, 12021 S. Wilmington Ave., Los Angeles, CA 90059; phone: (310) 668-4617; fax: (310) 223-0740; e-mail: Doogunye@cdrewu.edu

INTRODUCTION

Passing scores on the U.S. Medical Licensure Examination (USMLE) Steps 1 and 2 are required at the end of medical school and eventually for physician licensure.[1] Residency programs critically evaluate USMLE performance for residency placements.[1] The identification of predictors of students at risk of failing the USMLE examinations would be beneficial to student educators and students themselves, providing the opportunity for early intervention. Armstrong,[2] Myles[3] and Ogunyemi[4] have all noted a positive correlation between USMLE 1 and end-of-clerkship National Board Medical Examiners shelf-examination in Obstetrics and Gynecology (NBME-OB/GYN) examination.

A study by Myles[5] showed a correlation between USMLE Steps 1 and 2. Students failing either the USMLE Step 1 or NBME-OB/GYN were more likely to fail the USMLE Step 2. Demographic factors, such as race, age and gender, may have significant correlations with students' learning abilities. Undergraduate Grade Point Average (GPA) and Medical College Admission Test (MCAT) scores used by medical colleges in determining eligibility for college admission can potentially predict at-risk students. Since faculty train, work with and observe students, it is likely that faculty assessment of students may correlate with students' performance on standardized examinations. We could not find any previous study that evaluated most of these possible predictive factors of USMLE Step-2 examination in the same cohort of medical students. Thus, we sought to evaluate all these factors in a medical student population attending a diverse U.S. medical school.

The purpose of this study was to assess the correlation between USMLE Step-1 scores, undergradu-

ate GPA, MCAT scores, NBME-OB/GYN, faculty evaluations, demographic factors and the USMLE Step-2 scores.

The hypothesis was standardized examinations, faculty assessments, and demographic factors would correlate with USMLE Step 2 results.

Material and Methods

Charles Drew University, College of Medicine is a medical school with a diverse population located in the low-income urban area of South Central Los Angeles. Each class is small, with an average of 24 students per class. The university has a focus on primary care. The students are involved in a longitudinal primary care program that includes a continuity clinic in the third year and a primary care research project requiring a report and presentation in the fourth year.

The students are integrated into the obstetrics and gynecology service, working intimately with the faculty and residents. They receive faculty lectures that cover the educational objectives outlined by the Association for Professors of Gynecology and Obstetrics (APGO).[6] Each medical student is graded by faculty and residents for their clinical evaluation at the end of the rotation using a seven-point Likert-type scale. There are eight categories—history taking, physical examination, oral case presentation, communication skills, fund of knowledge, clinical judgment, professional attitudes and behaviors, and procedural skills—on which students are graded. The scores are averaged to obtain an overall clinical competence score. A score of 1 is unsatisfactory and a score of 7 is superior. Students must achieve an overall clinical competence score of 4 or more as part of the criteria for passing the OB/GYN clinical rotation. The clerkship director does not participate in the clinical evaluation. The evaluating faculty and residents do not have access to the MCAT scores, USMLE Step-1 scores nor the NBME-OB/GYN results. The students' final grade for the OB/GYN rotation is compiled based on their clinical performance, an oral examination, a Microsoft® PowerPoint® presentation to faculty and residents, and their score on the NBME-OB/GYN examination.

This was a retrospective records review. Institutional Review Board (IRB) approval was obtained. The NBME-OB/GYN scores were obtained from departmental records. USMLE scores, MCAT scores, undergraduate overall and science GPA, and demographic data were obtained from the Student Affairs Office. The records of 171 medical students who rotated through the OB/GYN department from 1992–2001 formed the basis of this study. There were 219 total records; however, 48 records were incomplete and excluded from the analysis. Pearson correlation and Spearman correlation coefficients were used to assess the correlations between USMLE scores and NBME-OB/GYN scores, race, age, sex, MCAT and GPA scores. To assess variables that were independently associated with a failing score on the USMLE Step-2 examination, a logistic binary regression analysis was performed by entering all the variables that had significant correlations ($p<0.05$) into the model. The variables entered included undergraduate overall GPA, science GPA, USMLE Step-1 score, number of MCAT attempts, NBME-OB/GYN score, students' age during the third-year OB/GYN rotation, academic year of the third-year rotation and faculty evaluation grades. All the variables were divided into two groups based on the passing score; for USMLE Steps 1 and 2, the passing score nationally is 182; for undergraduate GPA and science GPA, the selected grade was 3 (B average); for faculty evaluation grade, the passing grade was 4 or more; and for the NBME-OB/GYN exam 11th percentile was used as a passing score. The latter two criteria were set by the department of OB/GYN. The students were divided into two groups based on their age; those greater than and less than the age of 30 years. Paired Students' t test and ANOVA tests were used as appropriate. P values of <0.05 were taken as significant.

Results

The mean age of the students during the third year OBGYN clerkship was 30.26 (standard error of the mean ± 0.34) years. Forty-six percent were male and 54% were female. Forty-six percent were African-American, 35% were Hispanic, 12% were

Table 1. Demographics and examination scores of medical students (N=171)

Variable	Value
Male medical students	78 {46%}
Female medical students	93 {54%}
Caucasian medical students	12 {7%}
African-American medical students	79 {46%}
Hispanic medical students	59 {35%}
Asian medical students	21 {12%}
Mean students age in third year	30.26
Mean NBME-OB/GYN Score	67.47 (7.7)
Mean NBME-OB/GYN Percentile	40.57 (2.3)
Mean MCAT score	24.8 (5.03)
Mean undergraduate science GPA	3 (0.46)
Mean undergraduate overall GPA	3.1 (0.40)
Mean USMLE Step-1 score	194.53 (23.44)
Mean USMLE Step-2 score	190.63 (25.04)
Mean faculty evaluation grade	5.05 (1.01)

{} Percentage; () Standard error of the mean; NBME: National Board of Medical Examiners; USMLE: U.S. Medical Licensure Examination; MCAT: medical college admission test; GPA: college grade-point average

Asian, and 7% were Caucasian (Table 1). The mean NBME-OB/GYN score and percentile, respectively, were 67.47 (±7.7) and 40.57 (±2.3). The mean USMLE Step-1 and Step-2 scores were 194.53 (±23.44) and 190.63 (±25.04). The mean MCAT score was 24.8 (±5.03), and the mean GPA was 3.1 (±0.4).

Table 2 shows the correlation coefficient and P value between USMLE Step 2 and other variables. Significant correlations were found between the USMLE Step-2 exam scores and undergraduate GPA (r=0.287, p=0.000), MCAT score (r=0.524, p=0.000), USMLE Step-1 score (r=0.681, p=0.000), advancing year of clerkship rotation (r=0.432, p=0.000), NBME-OB/GYN score (r=0.614, p=0.000), increasing age of the student at the time of the third-year clerkship rotation (r=-0.405, p=0.000), and faculty evaluation grade (r=0.400, p=0.000). There was no correlation between race and gender with the USMLE Step-2 scores.

The variables that were independently associated with a failing score on the USMLE Step-2 examination are shown in Table 3. A failing score on the NBME-OB/GYN was the most independently significant variable for a failing score on the USMLE Step-2 examination (p=0.008), followed by a USMLE Step-1 failing score (p=0.01), faculty evaluation grade of 4 or less (p=0.029) and multiple attempts on the MCAT examination (p=0.033). Undergraduate GPA, academic year of third-year rotation and students' age did not independently correlate with a failing USMLE Step-2 score.

Discussion

Our findings show that there is a significant independent correlation between the USMLE Step-2 score and the NBME-OB/GYN score. Students who failed the NBME-OB/GYN examination were five times more likely to fail the USMLE Step 2 on the first attempt. The NBME-OB/GYN shelf examination as an example of a core clinical rotation highlights the significance of clinical training on USMLE Step-2 results. This is further supported by the significant negative correlation between the number of attempts on the NBME-OB/GYN shelf examination and the USMLE Step-2 score. As the USMLE Step 2 evaluates the overall third-year clinical training, it is reasonable to expect some association between the NBME-OB/GYN score and the USMLE Step-2 score. Thus, the NBME-OB/GYN shelf examination may be useful in selecting students in need of remediation before taking the USMLE Step-2 examination. It also suggests that using the standardized NBME-OB/GYN shelf examination as the end of clerkship evaluation rather than an in-house examination maybe beneficial. This is supported by a review of in-house examinations by Jozefowiz et al., which showed that most were of relatively low quality and substantially different from standardized national examinations.[7] These findings also confirm those of Myles.[4]

The USMLE Step-1 score also showed a strong but lesser independent correlation with USMLE Step-2 score. A student failing the USMLE Step 1 was 3.8 times more likely to fail the USMLE Step 2. Perhaps this lesser correlation is due to the differing focus of the USMLE Step-1 and Step-2 examinations. The USMLE Step 1 tests basic science knowledge and pathophysiology, while the NBME-OB/GYN and USMLE Step-2 examinations test the application of these principles based on the ability to obtain a pertinent history, clinical assessment and develop management plan. The linear correlation of the two standardized tests with USMLE Step-2 scores suggests a continuum of understanding and development of analytical skills from preclinical to

Table 2. Correlation between USMLE Step-2 scores and other variables

Variable	Rho Correlation Coefficient	P Value
Year of third clinical year rotation	0.432	0.000
Undergraduate science GPA	0.255	0.000
Undergraduate overall GPA	0.287	0.000
MCAT score	0.524	0.000
USMLE Step-1 score	0.681	0.000
NBME-OB/GYN score	0.614	0.000
Faculty evaluation grade	0.400	0.000
African-American race	-0.050	ns
Female gender	0.044	ns
Increasing age during third year clinical rotations	-0.405	0.000
Increasing number of MCAT attempts	-0.182	0.000
Increasing number of NBME-OB/GYN exam attempts	-0.310	0.000

NBME: National Board of Medical Examiners; USMLE: U.S. Medical Licensure Examination; MCAT: medical college admission test; GPA: college grade-point average; OB/GYN: obstetrics and gynecology

clinical instruction.

Undergraduate GPA and MCAT scores showed a positive linear correlation with students' performance on USMLE Step 2, but when analyzed as independent variables, they were found not to be predictive. However, students who had multiple attempts on the MCAT examination were 2.8 times more likely to fail the USMLE Step 2 on the first attempt. These findings suggest that poor standardized test takers may have similar difficulties with other standardized tests regardless of the subject matter.

Faculty evaluation grades were used as a measure of the quality of clerkship instruction. The faculty evaluations are completely independent of the standardized tests since none of the faculty have access to any of the students' educational records and have no contact with the Student Affairs Office. The positive correlation between students' grades and faculty evaluation could be because faculty have to assess the fund of knowledge of students, their professionalism and clinical judgment. Thus, students who receive high grades have been perceived by faculty to have learned the objectives of the clerkship and vice versa for those who got poor grades. Students who got a low faculty evaluation grade were 3.4 times more likely to fail the USMLE Step-2 examinations on the first attempt. This suggests that subjective assessment by faculty could be an objective risk assessment of failure on the USMLE Step 2.

In this diverse group of medical students, it was interesting to note that race and gender had no significant correlation with USMLE Step-2 scores. Increasing students' age correlated with decreasing USMLE scores but was not significant on logistic regression analysis.

The limitations of this study include the retrospective design. These results may not be applicable to other institutions due to the racial/ethnic make-up of this study population. From 1992 to 2001, 664,631 students were enrolled in American medical schools.[8] Of those students, 66% identified themselves as Caucasian, 18% Asian, 7.6% African American and 2.5% Hispanic.[8] The racial/ethnic make-up of this study population over the same period of time was 46% African-American, 35% Hispanic, 12% Asian and 7% Caucasian.

In summary, failing the USMLE Step-1 or NBME-OB/GYN examinations increased the risk of failing the USMLE Step-2 examination. Faculty evaluation also detected students in danger of failing USMLE Step 2. These findings may allow for detection of at risk students for early intervention and

Table 3. Logistic regression analysis with a failing score on the USMLE Step-2 (<183) score as the dependent variable

Variables	P Value	Odds Ratio	95% CI for Odds Ratio	
			Lower	Upper
USMLE Step-1 score >183*		1.00		
USMLE Step-1 score ≤182	0.010	3.696	1.375	9.934
NBME-OB/GYN score percentile ≥11th*		1.00		
NBME-OB/GYN score percentile ≤10th	0.008	4.962	1.512	16.290
Single attempt on MCAT examination*		1.00		
Multiple attempts on MCAT examination	0.033	2.828	1.088	7.350
Faculty evaluation grade >4*		1.00		
Faculty evaluation grade ≤4	0.029	3.372	1.136	10.008
Undergraduate Overall GPA ≥3*		1.00		
Undergraduate Overall GPA <3	0.324	0.513	0.136	1.930
Undergraduate Science GPA ≥3*		1.00		
Undergraduate Science GPA <3	0.702	1.301	0.338	5.012
Students age ≤30 years*		1.00		
Students age >30 years	0.768	1.210	0.341	4.291
Year of third clinical year rotation ≥1997*		1.00		
Year of third clinical year rotation ≤1996	0.093	2.132	0.882	5.149

* reference variable (odds ratio=1.00); NBME: National Board of Medical Examiners; USMLE: U.S. Medical Licensure Examination; MCAT: medical college admission test; OB/GYN: obstetrics and gynecology; GPA: college grade-point average

remediation prior to sitting for the USMLE Step-2 examination.

REFERENCES

1. Berner ES, Brooks M, Erdmann IV JB. Use of the USMLE to select residents. *Acad Med.* 1993,68:753-755

2. Armstrong A, Dahl C, Haffner W. Predictors of performance on the National Board of Medical Examiners Obstetrics and Gynecology subject examination. *Obstet Gynecol.* 1998; 91:1021-2.

3. Myles TD. U.S. Medical licensure examination Step-1 scores and Obstetrics and Gynecology clerkship final examination. *Obstet Gynecol.* 2002;100(5 Pt 1):955-958.

4. Ogunyemi D, Taylor-Harris D. NBME Obstetrics and Gynecology Clerkship Final Examination Scores: Predictive Value of Standardized Tests and Demographic Factors. *Jour Reprod Med.* 2004;49:978-982.

5. Myles TD, Henderson RC. Medical licensure examination scores: relationship to obstetrics and gynecology examination scores. *Obstet Gynecol.* 2002;100(5 Pt 1):955-958.

6. Association for Professors of Gynecology and Obstetrics. Instructional objective for a curriculum in obstetrics and Gynecology. 7th ed. Washington, DC: APGO, 1997.

7. Jozefowicz RF, Koeppen BM, Case S, et al. The Quality of In-House Medical School Examinations. *Acad Med.* 2002:156-161.

8. Total Enrollment by Gender and Race/Ethnicity, 1992–2001. AAMC Data Warehouse: student section. January 16, 2002. ■

We Welcome Your Comments

The *Journal of the National Medical Association* welcomes your Letters to the Editor about articles that appear in the *JNMA* or issues relevant to minority healthcare. Address correspondence to ktaylor@nmanet.org.